AD-A219 626

Final Technical Report

Air Force Office of Scientific Research
AFOSR-87-0300

November, 1989

INDIANENSIS UNIVERSITATIS SIGILLUM
LUX ET VERITAS
MDCCCXX

Hearing and Communication Laboratory

Department of Speech and Hearing Sciences
Indiana University
Bloomington, Indiana 47405

SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 90-0801
6a. NAME OF PERFORMING ORGANIZATION: Charles S. Watson, Dept. of Speech & Hearing, Indiana Univ. — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NL
6c. ADDRESS (City, State, and ZIP Code): Bloomington, IN 47405	7b. ADDRESS (City, State, and ZIP Code): Building 410, Bolling AFB, DC 20332
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Air Force Office of Scientific Research — 8b. OFFICE SYMBOL (If applicable): NL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-87-0300
8c. ADDRESS (City, State, and ZIP Code): Bolling Air Force Base, DC 20332-6448	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2313; TASK NO. A6; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Perception of Complex Auditory Patterns

12. PERSONAL AUTHOR(S)

Watson, Charles S. (Indiana University), Kidd, Gary R. (Indiana University)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 9/1/87 TO 9/30/89	1989, November	19

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	complex sounds, auditory patterns, psychoacoustics

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

See other side.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: JOHN F TANGNEY	22b. TELEPHONE (Include Area Code): (202) 767-5021 — 22c. OFFICE SYMBOL: NL

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

This report describes the results of research in four areas: (1) the perception of complex sounds, including tonal sequences, multidimensional complex sounds, and Gaussian noise; (2) information integration; (3) multi-stage decision making; (4) studies of the relation between auditory abilities measured with speech and nonspeech stimuli. Major accomplishments during this funding period include: (a) the discovery of the proportion-of-the-total-duration (PTD) principle: Each individual component of a complex sound is resolved with an accuracy that is a function of its proportion of the total duration of the sound; (b) the demonstration that even small degrees of logarithmic frequency transposition of tonal patterns severely degrades the detectability of pattern changes in novel sequences, but not in familiar sequences; (c) advances in research showing that temporal integration of auditory information is limited by two distinct types of internal noise, one that is added at the periphery before a decision statistic is formed, and "central," or post-decision, noise; (d) the development of a theory that incorporates these limiting factors; (e) the completion and publication of studies of categorical perception for speech and non-speech sounds, demonstrating that enhanced discrimination performance in the region of certain categorical boundaries does not, as previously theorized, reflect either "hard-wired" feature detectors in the auditory nervous system, nor psychoacoustic boundaries determined by acoustic peculiarities of complex waveforms.

Final Technical Report

Air Force Office of Scientific Research
AFOSR-87-0300

November, 1989

The following technical report describes work accomplished between 1 September 1987 and 30 September 1989. Work completed prior to this period was described in the original proposal for the current award and the HCL Progress Reports No.s 1, 2, and 3 which which have been previously circulated.

ABSTRACT

This report describes the results of research in four areas: (1) the perception of complex sounds, including tonal sequences, multidimensional complex sounds, and Gaussian noise; (2) information integration; (3) multi-stage decision making; (4) studies of the relation between auditory abilities measured with speech and nonspeech stimuli. Major accomplishments during this funding period include: (a) the discovery of the proportion-of-the-total-duration (PTD) principle: Each individual component of a complex sound is resolved with an accuracy that is a function of its proportion of the total duration of the sound; (b) the demonstration that even small degrees of logarithmic frequency transposition of tonal patterns severely degrades the detectability of pattern changes in novel sequences, but not in familiar sequences; (c) advances in research showing that temporal integration of auditory information is limited by two distinct types of internal noise, one that is added at the periphery before a decision statistic is formed, and "central", or post-decision, noise; (d) the development of a theory that incorporates these limiting factors; (e) the completion and publication of studies of categorical perception for speech and non-speech sounds, demonstrating that enhanced discrimination performance in the region of certain categorical boundaries does not, as previously theorized, reflect either "hard-wired" feature detectors in the auditory nervous system, nor psychoacoustic boundaries determined by acoustic peculiarities of complex waveforms.

Scientific Goals

The goals of the project were essentially unchanged from those described in the original proposal. We continued to pursue several lines of experimentation with complex (non-speech) auditory patterns and also extended this work in lines of research begun when the current grant was awarded in September, 1987: (a) studies of immediate auditory memory as a limiting factor in the discrimination of complex auditory patterns; (b) replication of our pattern discrimination studies with real and synthetic speech stimuli; and (c) multi-dimensional pattern processing.

Research Summaries

I. Complex Auditory-Pattern Perception

A. Tonal Sequences

1. *Proportional target-tone duration as a factor in auditory pattern discrimination.* C. Watson, Kidd, and Washburne

Results of previous experiments with isochronous tonal-patterns have appeared to reflect an informational limit on pattern processing, a limit that appeared to be reasonably close to the "magical" 7+/-2 (components in discriminable patterns). However, these results are also consistent with the alternative hypothesis, that target tones are equally well resolved if they occupy equal proportional durations of the patterns in which they occur. The results of a new series of experiments revealed that the target-tone's proportion of the total pattern duration accounts for most of the variance, while n, total duration, and individual component duration account for relatively little. These results were obtained for patterns ranging in duration from 100 msec to 2 sec. Additional experiments were designed to test the limits and generality of the proportion-of-the-total-duration (PTD) rule. These experiments have shown that 1) pattern discrimination is not significantly affected by the ratio of context to target tones, except when the context tones are very brief in absolute duration as well as relative duration, 2) discrimination is primarily determined by PTD for different target positions in spite of increases in target-position uncertainty, 3) PTD still accounts for discrimination performance when two non-adjacent target tones are used, but performance levels are slightly lower than in the single-target case for a given PTD. (A manuscript describing this work is in preparation.)

2. *Detection of pattern repetition in continuous tone patterns.* Watson, Kidd, Washburne.

Additional data have been collected in experiments designed as variations on those reported by Guttman and Julesz (1963, Journal of the Acoustical Society of America), but using tonal sequences rather than noise samples. Subjects are presented with repeating or non-repeating tone patterns, using a tracking paradigm that increases or decreases the number of tones in a pattern depending on a subject's ability to detect repetition. The data show predicted strong effects of tone duration and bandwidth, as well as a significant interaction (due to a slightly greater effect of bandwidth at the 50-msec tone duration). Listeners were able to detect the repetition of patterns consisting of more tones with the shorter tone duration and the wider bandwidth. Despite our attempts to minimize the occurrence of "unique events", subjects' reports indicated that judgments were often based on the recurrence of particular events rather than detection of whole-pattern repetition.

3. *Perception of Frequency- and Time-Transposed Auditory Patterns.* Kidd, C. Watson.

a. *Transposition in frequency and time.*

This series of experiments examined the effect of transposition in frequency and in time on the detection of single-tone frequency changes in unfamiliar tonal patterns. In the first experiments, listeners' abilities to discriminate between correct and incorrect transpositions of randomly generated

five-tone patterns were investigated. Frequency transposition resulted in very large decrements in performance, which occurred with the smallest amount of transposition, with a tendency for additional reduction in performance with increases in the amount of transposition. Time transposition (i.e., increases in tone duration) resulted in little or no performance decrements after 2-3 hours of training.

b. *Frequency transposition with smaller transposition ratios.*

In the frequency-transposition experiment described above, the smallest amount of transposition (a 1.15 frequency ratio) resulted in performance decrements that were nearly as great as the larger transposition intervals. A second experiment was run to determine whether there is a gradual or an abrupt change in performance for smaller transposition intervals. The frequency-transposition experiment was repeated using transposition ratios of 1.00, 1.03, 1.06, 1.09, 1.12, 1.15 (i.e., five log steps from 1.0 to the smallest interval used in the previous experiment). Performance was found to decrease gradually over this range for all but one of the four subjects who showed a rather abrupt decrease in performance.

c. *Transposition with reduced uncertainty.*

i. Reduced transposition uncertainty. To test the extent to which transposition performance is adversely affected by transposition uncertainty (i.e., uncertainty with respect to the amount of transposition on a given trial), the frequency-transposition experiment was run again with a single transposition ratio (1.52). Even after 10 hours of practice at this transposition ratio, little or no improvement in performance was obtained. Considerable trial-to-trial stimulus variation remained in this experiment, however, and may be responsible for some of the poor performance.

ii. Reduced pattern uncertainty. To test the extent to which transposition performance is affected by pattern uncertainty (or pattern familiarity), another frequency-transposition experiment was run in which the same pattern (randomly generated) was presented on every trial. Under these conditions, transposition performance improved dramatically; In many cases, transposition effects were completely absent.

d. *Frequency transposition with 2-tone patterns.*

A related line of investigation has been conducted using two-tone patterns (i.e., single intervals). These experiments were designed to provide a more direct test of the basic ability to detect changes in relative frequency by using the simplest possible stimuli required to test this ability. Performance under high-uncertainty conditions was similar to that for 5-tone patterns, although complex interactions among interval size, direction of transposition, and direction of frequency changes were observed. Results under low uncertainty conditions indicate that transposition effects and interaction are diminished and performance is often as accurate as in non-transposed conditions.

General Conclusions: Resolution of the pattern of relative frequency changes in transposed tonal sequences is severely degraded under conditions of high stimulus uncertainty for patterns of five or more components. Resolution of relative frequencies within transposed patterns is nearly as accurate as for non-transposed patterns a) under conditions of minimal stimulus uncertainty, or b) when the patterns contain only two components. These effects of stimulus uncertainty and complexity are the

same as those previously reported in pattern discrimination experiments without frequency transposition.

4. *Discrimination of multiple-tone targets in tonal sequences.*

We have completed an initial investigation of the ability to discriminate and identify multiple-tone targets embedded within tonal sequences. This work extends earlier experiments with single-tone targets to cases in which the target itself is a pattern. In this initial experiment, one of four three-tone targets is presented on each trial in the middle of a nine-tone sequence in which all tone durations are set at 100 ms. A same/different task is used in which listeners must determine whether the three-tone target in the temporal center of the pattern is the same in both patterns presented on a trial. The six context tones are selected randomly for each pattern on every trial. Thus, the two patterns on each trial are always different, but the three-tone target segment may be the same. On "different" trials the target tones are permuted (thus preserving the distribution of spectral energy). The level of the context tones is initially set 40 dB down from the target-tone level (75 dB SPL) so that target tones can be easily "heard out". An adaptive tracking procedure is used to adjust the level of the context tones according to listeners' ability to detect the target sequence. The results have revealed large differences across listeners and patterns, but in the best cases the level of the context tones is increased from roughly 40-to-50 dB to 0-to-10 dB below the target tones over the course of 150 hours of training (over 6000 total trials). In the worst cases there is little or no improvement even after 150 hours of training. (Dr. Kathy Barsz was a visiting scientist during July and August, 1989, and participated in the design of this experiment.)

Results with lower levels of uncertainty (using a single target sequence instead of four) have shown no additional improvement compared to ending levels under higher uncertainty. However, the performance of a new listener with no prior experience in the high-uncertainty experiment suggests that learning occurs at a faster rate with lower uncertainty. Additional experiments using patterns with other component and total durations will be conducted to examine the role of those variables in the detection of complex multi-component targets.

5. *Perception of salient auditory events or figures.* C. Watson, Kidd, Washburne.

We continued to use our auditory figure-identification procedure to study factors that may be systematically related to the emergence of auditory figures or "targets" from various backgrounds.

6. *Perception of complex auditory patterns by humans.* Espinoza-Varas, C. Watson.

A chapter, writtten by B. Espinoza-Varas and C.S. Watson, was contributed to "The comparative psychoacoustic of complex acoustic perception," co-edited by S. Hulse and R. Dooling (in press, Lawrence Earlbaum Associates, N.J.). The thesis proposed is that, for the case of mammals, measured phylogenetic differences in pattern perception ability appear to result more from differences in information processing abilities (i.e., central factors), than from structural and/or physiological differences in the peripheral auditory system. (A review of this book commented favorably on the chapter by Espinoza-Varas and Watson (Science, December 1, 1989).)

B. Spectrally-Complex Sounds

1. *Perception of multidimensional complex sounds.* Kidd, C. Watson, Washburne.

This project extended our work on listeners' abilities to categorize, or classify, complex sounds, based on information in three independent acoustic "features", studying abilities to 1) allocate attention simultaneously to three complex auditory dimensions, and 2) accumulate complex multidimensional information over a series of sound pulses. The stimuli consisted of sequences 1, 3, 5, or 7 100-msec sound pulses. The features were spectral shape, temporal envelope, and harmonicity of generating sinusoids. Listeners can integrate information in sequences of sound pulses, with little or no loss of efficiency with increasing sequence length. Although listeners' absolute efficiency in this task is fairly high, comparable performance levels are achieved in spite of a variety of individual patterns of attention to the features of the stimuli. Attempts to train listeners to attend to features to which they initially do not attend met with little or no success.

C. Gaussian-noise Stimuli

1. *Discriminability of noise samples.* Robinson, Fallon, Rickert.

Our work on the discriminability of noise samples was continued. This work employs pairs of complex auditory waveforms that listeners are asked to discriminate. The waveforms are samples of broad-band, white, Gaussian noise. All experiments made use of a same-different paradigm. Research reported in previous years has investigated several parameters that affect the discriminability of noise samples. In our initial work in this area, we demonstrated that if only a small segment of a noise burst is replaced with a new sample of noise, the temporal position of the new segment is of critical importance in determining discriminability. Independent of total burst duration, changes made at the end of a burst are more discriminable than those in the middle, which are more discriminable than changes made at the beginning. In that work, we also observed a 'Weber's law' relationship between the duration of the altered segment and total burst duration. As total burst duration was varied from 25-msec to 150-msec, the ratio of the duration of the altered segment to total burst duration remained constant within each of the three conditions: beginning, middle, and end. In two additional experiments, we investigated the effects of decorrelating the noise samples on "different" trials. This was done in two ways: first, by time delaying the noise presented in the second interval relative to that in the first interval and, second, by adding an independent sample of noise to the sample presented in the first interval to generate the sample presented in the second interval. In both cases, increasing the correlation of the samples in the two intervals reduces discriminability. However, the functions relating performance to correlation are different depending on the method used to vary the correlation. In other work, we have attempted to improve performance by reducing the size of the set of noise samples presented to the subjects. We found a negligible effect of set size. We have previously reported on the effects of introducing temporal gaps in the noise bursts. In the last year, we have extended the work on temporal gaps and have investigated the effects of reducing the level of portions of the noise bursts.

a. *Effect of gap duration and position.* Robinson, Fallon.

In this experiment, the overall duration of the bursts of noise in a pair was 150 msec and either a 25-msec segment of new noise was inserted at the end of the burst or a 50 msec sample was

appended to the beginning of the burst. A silent interval or gap replaced a portion of the repeated segment either immediately following or immediately prior to the appended segment. The duration of the gap was gradually increased until only 5-msec of the repeated or inserted segment remained. Although discriminability increased as gap duration increased, the presence of a brief repeated segment temporally separated from the appended segment by 90-120 msec caused a large decrement in performance. For example, when each burst in a pair consisted of a 5-msec repeated segment followed by a 120-msec gap and a 25-msec appended segment, the average P(C) is 0.72. If the 5-msec repeated segment was not present and the pair of 25-msec bursts was presented in isolation the overall P(C) increased to 0.88. It would appear, then, that interactions occurring after such a long silent interval are unlikely to be due to peripheral sensory interactions, as was suggested by Hanna (*Perception & Psychophysics*, 36, 409-416, 1984).

b. Effect of varying the level of the repeated segment. Robinson, Fallon.

The effects of decreasing the level of the repeated segment on the discriminability of pairs of 150-msec bursts of noise was examined in two conditions. In the first condition a 15-msec sample of new noise was appended to the end of the second sample of pairs presented during "different" trials. A 75-msec sample of new noise was inserted at the beginning of the second sample of a pair in the second condition. Prior to varying the level of the repeated segment, the discriminability of the bursts presented in the two conditions was similar. Discriminability tended to gradually increase as the level of the repeated segment was decreased. Performance did not greatly improve until the level of the repeated segment was reduced from 50 to at least 28 dB SPL, suggesting that peripheral masking cannot account for the effect of the temporal position of the appended or inserted segment of noise on discriminability.

c. Discriminability of Noise Burst under Dichotic Listening Conditions. D. Robinson and M. Rickert.

Several binaural phenomena (e.g., the precedence effect) suggest that under some conditions interaural differences occurring at the beginning of a brief acoustic event are more effective cues than those occurring later. In the context of our previous work concerning the discriminability of diotic noise samples, these phenomena suggest that the advantage of the 'end' condition might not be found with dichotic presentations. Thus one goal of extending our research to include binaural listening tasks was to determine whether the factors affecting performance for diotic presentations act in a similar fashion under dichotic conditions. In this experiment, performance was measured using a single-interval, yes-no procedure. The subject's task was to detect the presence of an interaurally uncorrelated segment of noise within a binaural broadband noise burst. Three overall stimulus durations (25, 75, and 150 ms) were used. Psychometric functions were then obtained by systematically varying the duration of the segment of uncorrelated noise at each of three temporal positions (beginning, middle, and end). The results can be summarized as follows. First, independent of overall stimulus duration, detectability improves as the duration of the uncorrelated segment is increased; percent correct is a monotonic increasing function of the duration of uncorrelated noise. Second, for most durations of uncorrelated noise, detectability is best when the segment occurs at the end of a burst. Relative to diotic presentations, however, the performance decrements obtained by varying the temporal position of the uncorrelated segment are much smaller. This result is not easily accounted for by most models of binaural processing. For example, the EC

model [Durlach, N.I. J. Acoust. Soc. Am., 35, 1206–1218 (1963)] assumes that performance is determined by the power of the difference in the waveforms in the two auditory channels. Since power is computed over the total stimulus duration, it is independent of the temporal position of the uncorrelated noise. Third, for each stimulus combination used in this study, performance for dichotic presentations greatly exceeds performance for diotic presentations.

d. *Molecular Psychophysics.* Robinson, Fallon

In their work concerning the role of stimulus uncertainty in the discrimination of complex tonal patterns, Watson, Kelly and Wroton (1976) and Watson and Kelly (1981) distinguish between central and peripheral factors that limit the processing of sensory information. They argue that the limits on information processing that can modified by manipulating stimulus uncertainty (i.e., smaller set size, over-training, etc.) are central, whereas those that cannot be modified are peripheral. The main goal of the research described here is to determine whether the discriminability of different 'target' segments (i.e., beginning, middle, and end) within a random, time-varying waveform can be altered by minimizing stimulus uncertainty.

Previously, we have used free-running samples of noise such that new waveforms were generated prior to each trial. This method made it impossible for subjects' to base their decisions on long-term information about the spectral-temporal characteristics of the noise samples. If, however, the factors affecting discriminability of deterministic tonal sequences and stochastic waveforms are similar, then it may be possible to eliminate the effects of temporal position by (1) reducing the size of the stimulus set from which samples are drawn and by (2) over-training.

Two separate experiments, both employing a two-interval, same-different method were conducted. All stimuli used were samples of computer-generated broad-band Gaussian noise, 150 msec in total duration. The first experiment was designed to estimate the hit and false alarm rates for each of twenty 'Same' pairs and twenty'Different' pairs. This experiment measured subjects' relative performance for each of the individual samples. The data indicate that all pairs of noise samples are not equally discriminable. There is considerable scatter when the points are plotted in ROC space. Although these samples have identical long-term statistical properties, it is likely that the observed variability among samples is due to differences in their short-term power or energy spectra. Thus some pairs of noise samples are more discriminable than others. A subset of these samples was then used in the second experiment. The samples which were selected were chosen to maximize their separation in ROC space. In this experiment only two noise bursts (one 'Same' and one 'Different') were run in each experimental block. Moreover, the same noise bursts were used in successive blocks. Several experimental sessions were conducted for each pair of noise bursts in order to monitor changes in discrimination performance that might have resulted from over-training. We were unable to find evidence for improved performance with this reduced stimulus set and a large number of trials.

D. Information integration: Multiple observations and internal noise. Robinson, Berg.

The work described in this section began with two major goals. The first was to understand the processes by which humans integrate information over time or over channels. The "multiple look"

problem is the basis for our initial work in this area. The basic question is, "How much additional information is gained by allowing observers more than one observation in a detection or discrimination task?" The second goal was to develop and evaluate models of "internal noise." The amount and rate of improvement in performance with an increasing number of observations will depend not only upon the amount of internal noise, but upon the level of processing at which the internal noise is added.

Previous research (Swets, Shipley, McKey, and Green, 1961; Swets and Birdsall, 1978; Swets, Green, Getty, and Swets, 1978) has demonstrated that performance in signal-in-noise detection tasks improves as listeners are allowed more observations. According to signal detection theory, d' should increase as the square-root of the number of observations. This result assumes that the decision statistic is any monotonic transformation of the mean of the *n* likelihood ratios obtained from the *n* observations. In some versions of the derivation, internal noise is added prior to the formation of the decision statistic in order to account for the less than optimal performance at $n=1$. This does not alter the square-root-of-*n* prediction. Previous research, usually involving the detection of repeated signals in a noise background, has supported the square-root-of-*n* prediction. However, the earlier work provided only limited tests, since the use of signals in a noise background precluded exact specification of the noise and signal-plus-noise distributions on an observer's decision axis. Further, only small values of *n* were used. Our research has extended this work by developing a unique paradigm that allows exact specification of the distributions on an observer's decision axis (Berg and Robinson, 1987). In this paradigm, the observer's task is to judge from which of two distributions on frequency a sequence of *n* tones was sampled. Results from several experiments indicate that listeners can approach the theoretical d' for $n=1$, but do not follow the square-root-of-*n* rule, even for small *n*.

In order to account for the failure of the square-root-of-*n* prediction, we have assumed an additional source of internal noise (Robinson and Berg, 1986; Berg and Robinson, 1987; Berg, 1987). This type of internal noise may be thought of as additional variance introduced by uncertainty of the decision criterion, changes in response bias, or memorial factors associated with the decision statistic. Briefly, then, in the "partitioned variance" model, as we have termed it, internal noise is added at two stages: (1) at the periphery, before a decision statistic is formed and (2) centrally, after the statistic is formed. As we have previously reported, the model does an excellent job of describing the data of the sequential tone experiments, as well as those of a similar experiment in which the tones are presented simultaneously (Grantham, 1987). We have also developed techniques for estimating the amount of internal noise added at each of the two stages and techniques for determining the degree to which each tone in the sequence contributes to the observers' final decision (Sorkin, Robinson, and Berg, 1987). Our results indicate that subjects are highly idiosyncratic in the degree to which various portions of a temporal sequence contribute to their decision: some subjects appear to show a primacy-recency effect while others show roughly equal weighting of the tones in a sequence.

1. *Distribution of internal noise over the tonal sequence.*

An important question raised by our general model is whether information from each tone in a tonal sequence is equally weighted in determining an observer's decision. We have developed a technique for assessing how internal noise is distributed over the n-elements of a tonal sequence. In terms of the model, the amount of information obtained from different tones in an n-element

sequence will be reflected in the variance of the internal noise added at each temporal position. If a particular temporal position contributes little to the final decision, that position will be found to have a large amount of internal noise associated with it. If, on the other hand, a particular element contributes a great deal, that element will have less internal noise associated with it. Data from an auditory experiment were analyzed to assess how internal noise is distributed over successive temporal positions. Over many thousands of trials we store the frequency of the tones actually presented in the ith temporal position (i = 1, 2, ... n; where n is the number of tones in the sequence). We then partition these stored frequencies into bins of arbitrary width. The purpose of this analysis is to keep track of the number of trials on which the frequency of the ith element was in each frequency bin. For each bin and each temporal position, we then compute the probability that the subject responds that the sequence came from the lower distribution. Cumulative normal distributions are then fit to the resulting ogives. The standard deviation of the best fitting normal distribution is then an estimate of the standard deviation of the total internal noise limiting performance at each display position. Results have revealed a tendency for tones near the beginning and end of the sequence to contribute more to the final decision than tones in the middle.

E. Multi-stage decision making. Robinson.

In this work we considered a person-machine system consisting of an automated alarm and a human monitor. The task of the human is to monitor a noisy channel on which information about a potentially dangerous condition may appear. The alarm system monitors an independently noisy channel for information about the same threatening condition. Using basic concepts of statistical decision theory, the Contingent Criterion Model of such a person-machine system has been developed. According to the model, the human should establish two criteria for responding: one contingent on an alarm from the automated detector and one, on no-alarm. The model shows large gains in performance compared to either detector alone. In Progress Report No. 1, we reported experimental support for the model. In the last two years, we have continued to develop the Contingent Criterion model, to investigate the effects of additional variables on the performance of the model, and to expand the model to include target identification.

1. *Effects of inter-channel correlation.* In our initial development of the Contingent Criterion Model, we assumed that the noise in the alarm-system channel is uncorrelated with that in the channel monitored by the human operator. Such an assumption is probably unrealistic. We have recently investigated the degradation in performance that occurs with increased correlation between the two channels. The predictions of the model indicate that, although there may be a considerable performance decrement when the correlation is near unity, the model is quite robust, and system performance can exceed that of either detector alone even with correlations as high as 0.50.

2. *Effects of signal probability and values and costs.* Other parameters of importance in determining the performance of combined detection systems are the *a priori* probability of the signal and the values and costs of the possible outcomes of a decision. In our past work, we have evaluated system performance using measures derived directly from the Receiver Operating Characteristic (ROC), such as P(C) and d'. These measures are not affected by changes in *a priori* probability or

values and costs. Thus, we turned to measures based on Expected Value. Analyses based on Expected Value suggest that the advantage of a multi-stage decision making system over a single-stage one is highly dependent on *a priori* probability and values and costs. In fact, under some circumstances a single-stage system should be used.

3. *Target identification*. We expanded the Contingent Criterion Model to include signal classification (identification) as well as signal detection. This effort draws on the work of Nolte (JASA, 42, 773-777, 1967) Nolte and Jaarsma (JASA, 41, 497-505,1967), Green and Birdsall (Psych. Rev., 85, 192-206, 1978), and Starr, Metz, Lusted, and Goodenough (Radiology, 116, 533-538 ,1975). Our efforts to date suggest that the performance of a system consisting of a human operator and an automatic signal classifier can show significant improvement compared to either subsystem operating alone.

There are two important observations that may be drawn from this work. First, although combined system performance (human-plus-automated detectors) may be less than optimum with untrained operators, it is possible that human operators can be trained to use the available information more efficiently. A second observation is that system performance is dependent not only on the behavior (sensitivity and criterion placement) of the human operator and the criterion (threshold or alarm set-point) of the automated alarm system, but also on the *a priori* signal probability and the values and costs. System performance may be altered by changing any of these parameters of the system. We feel strongly that the designers of automated alarm systems must take into account the complex interactions which may occur when human operators are involved.

All of the work on multi-stage detection systems was done in collaboration with Dr. Robert D. Sorkin of Purdue University. We take this opportunity to again acknowledge his many contributions to this, and other projects, both at ARL and HCL. Portions of the work described here have been reported previously [Sorkin and Robinson (N.T.I.S. Report No. DOT/OST/p-34/85/021, 1985) and Robinson and Sorkin (In Trends in Ergonomics/Human Factors II, Eberts and Eberts, North-Holland: Elsevier, 1985)]. Additional support for this work was provided by contracts with the U. S. Department of Transportation and with the U. S. Naval Weapons Center, China Lake, CA.

F. Multiple Targets: Detection and identification of auditory signals in noise. Robinson, Rickert

The multiple target problem, and the consequent requirement for target categorization, has been simplified to the case of equally detectable, orthogonal signals. The set of potential signals used in this experiment consisted of four sinusoids (500, 1100, 1900, and 2700 Hz). In order to satisfy the condition of equal detectability across channels, the signal-to-noise ratio for each the four frequencies was adjusted to achieve a P(C) of approximately 0.75. We assumed that the wide frequency separation between signals was sufficient to meet the condition of orthogonality. That is, each signal frequency is processed independently. These assumptions allow application of a theorem relating target detection and identification (Starr, Metz, Lusted, and Goodenough, 1975). In a combined detection-identification task, subjects were asked to listen during a specified observation interval and to make two decisions: (1) report the presence or absence of a tonal signal in a broad-band gated

noise, and (2) indicate which one of the four signals occurred. The overall a priori probability of signal-plus-noise and noise-alone trials was balanced at 0.50. Further, each of the four frequencies was equally likely to occur on a signal-plus-noise trial. Group mean data indicate that detection performance exceeds identification performance by approximately 14 percent (0.77 to 0.63). Although more elaborate analyses are required to evaluate statistical fit to the data, the predictions made by the theorem are consistent with observed performance.

1. *Effect of changing the distribution of targets over channels.* Robinson, Rickert

We have completed a pilot study to investigate the effects of changing target probability over channels within the detection-recognition task. The experimental procedures described in the previous section were also used in this study. The a priori probability of a signal occurring was maintained at 0.50. However, on signal trials the probability of the 500 Hz signal was reduced to 0.10 while the probability of the remaining signals was increased to 0.30. Preliminary results indicate that identification performance increases relative to identification performance when the a priori probabilities of the four signals are equal.

II. Studies of the relation between auditory abilities measured with speech and nonspeech stimuli.

A. Categorical perception. Kewley-Port, C. Watson, Foyle

In March, 1988, we published a manuscript on the categorical perception of speech and non-speech sounds (Kewley-Port, Watson and Foyle, 1988). Conclusions were, 1) that a category boundary for these sounds was not obtained in tasks with reduced levels of stimulus uncertainty, 2) that this implies the categorization of familiar sounds is not a result of a psychoacoustic threshold, 3) that the functions obtained for the discrimination of temporal-onset differences for the speech and non-speech sounds were basically the same.

These conclusions were challenged by Richard Pastore in a Letter to the Editor of the *Journal of the Acoustical Society of America* (1988). A reply to Pastore's letter (Watson and Kewley-Port, in press) was published in the same issue. New data and analyses presented in our response further support the conclusions listed above, and directly contradict arguments raised by Pastore.

B. Detection thresholds for isolated vowels. Kewley-Port

A series of experiments on the detectability of vowels in isolation has been completed. Stimuli consisted of three sets of 10 vowels, one synthetic, one from a male talker and one from a female talker. Vowel durations ranged from 20 to 160 msec. Thresholds for detecting the vowels in isolation were obtained from well-trained, normal-hearing listeners using an adaptive-tracking paradigm. Detection thresholds for vowels calibrated for equal sound pressure at the earphones differed by as much as 20 dB across vowels. However, the same patterns of thresholds obtained for the vowel set were obtained for all vowel durations. An orderly decrease in vowel thresholds was obtained for increased duration, as predicted from temporal integration for pure tones. Several different analyses

have been conducted in an attempt to explain the differential detectability for the various vowels within each set. The most detailed of these analyses involved a model developed by Moore and Glasberg (1987) for calculating excitation patterns and the corresponding loudness level in phons. Although that model provided improved explanatory power over other models examined, further refinements of perceptual models of loudness will be needed to explain these data. (A manuscript describing these data has been submitted to the *Journal of the Acoustical Society of America.*)

C. Vowel Sequence Studies: High versus Minimal Uncertainty Detection using a Pattern Catalogue. Kewley-Port, C. Watson, Hackett.

Previous research in our lab has investigated complex auditory stimuli consisting of a series of brief tones presented one after the other to form single "word-length" tonal patterns. We have now replicated some of those experiments with speech stimuli. The first experiment examined the detectability of vowels, presented in sequences, under very high-uncertainty. Thresholds differed only slightly across temporal position in the sequence, but differed considerably for different subjects and vowel types. In a new study, a catalogue of 48 ten-vowel patterns was constructed. These stimuli were presented under high uncertainty, where all 48 patterns were randomly presented, and under minimal uncertainty where only a single pattern was presented. Results obtained for the detectability of vowels in the pattern catalogue under high uncertainty were similar to those obtained in the first experiment. Under minimal uncertainty, the detectability for vowels improved a great deal, but not as much as for tones in analogous tonal-pattern studies. Further experiments suggest that more energetic masking may be obtained for complex vowel spectra in vowel sequences than for pure tones. These experiments are now focusing on determining the parameters of vowel masking.

D. Effects of response uncertainty on the intelligibility of nonsense syllables. Espinoza-Varas, C. Watson, C. Parks.

Syllable intelligibility in the CUNY Nonsense Syllable Test (Dubno, J.R., and Levitt, H.J. (1981). *J. Acoust. Soc. Am.*, *69*, 249-261) was studied with a response format consisting of either the usual 7-9 alternatives or only 3- alternatives: the correct alternative plus the two most likely confusions. With a few exceptions, the average intelligibility of each of the syllables was consistently higher (about 15% on the average) in the 3-alternative condition, but there was a strong correlation between the performance in the two response formats.

E. Rate and number as limiting factors in the perception of sequences of syllables. C. Watson, Espinoza-Varas.

A test was developed to measure the effects of the number of syllables and, independently, those of syllable rate, on listener's abilities to process syllable sequences. The sequences consist of digitized and edited natural tokens of the spoken letters b, c, d, g, p, t, or v, produced as CV's ending in the vowel /i/. Results for an initial test group showed some individual's performance to be more limited by rate than by number, and vice versa. Some listeners show a much more severe drop when the number of syllables is increased from 2 to 4, than others. Basically, however, these "high capacity listeners"

(those who do not show a marked change in performance as n is increased from 2 to 4) are also the ones who do the best on the shortest sequences.

F. Informational limits in speech processing by normal and impaired listeners. Espinoza-Varas, C. Watson, Kyle.

Using a variation of the speech test described in the previous section, we compared the information capacity of normal and impaired listeners (moderate-to-severe losses). Percent information transmission, IT%, was estimated for each temporal position within the sequences, number of syllables, and syllable rate, for equal-sensation-level listening. For the normal listeners, 6 or more syllables could be identified, on the average, but for the best of the impaired listeners, the maximum number seems to be only 4-5. This would imply a loss of information capacity of as much as 30-40 percent. Additional Studies of this type are in progress.

G. Low communality between tests of auditory discrimination and of speech perception. Espinoza-Varas, C. Watson.

Factor analyses of the performance of normal listeners on the Test of Basic Auditory Capabilities, TBAC collected in a field (n=127) and using earphones (n=119) revealed similar structures for both groups, consisting essentially of two factors loaded on the nonspeech tests (complex sound/pitch and duration/intensity) and a third that is mainly a speech factor. In addition, multiple regression analyses showed that only a small portion of the variance in the speech tests can be accounted by the optimal linear combinations of the tonal test scores.

III. Partially Supported Projects

A. Response latency and decision criterion in psychophysical decisions. Espinoza-Varas, C. Watson, Patterson, Kyle.

A manuscript has been prepared for submission to *Perception and Psychophysics* describing earlier data on the effects of decision criterion on response latencies in psychophysical decisions. The response latency to a given stimulus was found to vary inversely with the distance between the stimulus and the current response criterion, and with the probability of the response. These effects were observed for both visually presented two-digit numbers and pure tones under three different decision tasks. (Original & experimental work supported by an NIH grant to the Central Institute for the Deaf, data analyses and manuscript preparation supported by NIH and AFOSR grants to Indiana University).

B. Development of norms and factor analysis of TBAC performance obtained in 398 listeners Espinoza-Varas, Watson

Since first developed (Watson et al., *J. Acoust. Soc. Am.* 1982, 71, S73), the eight-subset Test of Basic Auditory Capabilities (TBAC) has been administered to approximately 600 listeners, in 10

different studies carried out at the Boys Town National Institute, the University of Nebraska, Lincoln, University of Wisconsin, Madison, and the Speech and Hearing Center, Indiana University. This large data base offers a unique opportunity to standardize and develop norms for the TBAC and to perform factor and cluster analyses. We have recently consolidated the data from the various studies, which were previously stored in several different formats and media. We have developed a coherent structure for this data base and currently have TBAC results for 398 listeners entered in the CDC-856 Indiana University computer, which has available appropriate statistical packages. Preliminary analyses have confirmed that the TBAC is a reliable test, that the three-factor partitioning of the variance on the TBAC is a stable result, and that performance on these tests has an unexpectedly strong association with certain intellectual measures (WAIS scores, SAT's).

C. Relations between auditory capabilities and phoneme perception. Espinoza-Varas, C. Watson, Srygler.

A manuscript describing this work is in preparation for submission to the *Journal of the Acoustical Society of America*.

D. Interaural comparison of discrimination abilities in monaural hearing impaired listeners. Espinoza-Varas, C. Watson.

We continued to collect data to compare, interaurally, the discrimination abilities of monaural hearing impaired listeners. The goal is to assess the differential roles of central and peripheral processing in speech and non-speech tasks.

E. Articulatory-Rate context effects in phoneme identification. G. Kidd

These experiments demonstrated that the pattern of changes in articulatory rate over several syllables prior to a target phoneme can have a substantial influence on identification of the target phoneme. A manuscript describing this work was published in the *Journal of Experimental Psychology: Human Perception and Performance.*

Personnel

Name	Position Title	Department
Charles S. Watson, Ph.D.	Professor & Chair	Speech & Hearing Sci.
Donald E. Robinson, Ph.D.	Professor	Psychology
Diane Kewley-Port, Ph.D.	Assistant Professor	Speech & Hearing Sci.
Gary R. Kidd, Ph.D.	Assistant Scientist	Speech & Hearing Sci.
Blas Espinoza-Varas, Ph.D.	Visiting Assistant Professor	Speech & Hearing Sci.
Kathy Barsz, Ph.D.	Assistant Professor (Visiting Scientist)	Psychology SUNY Geneseo

Advanced Degrees Awarded

Bruce G. Berg, Ph.D., 1987; "Internal noise in auditory detection tasks."

Susan M. Falon, Ph.D., 1989; "Discriminability of Bursts of Reporducible Noise"

Manuscripts and Publications

Berg, B.G. (1987). Internal noise in auditory detection tasks. Ph.D. Dissertation, Indiana University.

Espinoza-Varas, B. (1987). Levels of representation of phonemes and bandwidth of spectral and temporal integration. In: M.E.H. Schouten (Ed.) The Psychophysics of Speech Perception, M. Nijhoff, The Netherlands, 80-90.

Espinoza-Varas, B. (1987). Involvement of the critical band in identification, perceived distance, and discrimination of vowels. In: M.E.H. Schouten (Ed.) The Psychophysics of Speech Perception, M. Nijhoff, The Netherlands, 306-313.

Sorkin, R.D., Robinson, D.E., & Berg, B.G. (1987). A detection theory method for the analysis of visual and auditory displays, Proceedings of the 31st Annual Meeting of the Human Factors Society, Vol. 2, 1184-1188.

Watson, C.S. (1987). Uncertainty, informational masking and the capacity of immediate auditory memory. In W.A. Yost and C.S. Watson (Eds.) *Auditory Processing of Complex Sounds*, Erlbaum Associates, Hillsdale, NJ.

Yost W.A. & Watson C.S. (Eds.) (1987). *Auditory Processing of Complex Sounds*, Erlbaum Associates, Hillsdale, NJ.

Kewley-Port, D., Watson, C.S., and Foyle, D. (1988). Auditory temporal acuity in relation to category boundaries: Speech and nonspeech stimuli. *J. Acoust. Soc. Am.*, *81*, 1133-1145.

Kidd, G.R., & Greenwald, A.G. (1988). Attention, rehearsal, and memory for serial patterns. *American Journal of Psychology*, *101*, 259-279.

Leek, M.R. & Watson, C.S. (1988). Auditory perceptual learning of tonal patterns. *Perception & Psychophysics*, *43 (4)*, 389-394.

Watson, C.S., and Kewley-Port, D. (1988). Some remarks on Pastore (1988). *J. Acoust. Soc. Am.*, *84*, 2266-2270.

Espinoza-Varas, B. and Watson, C.S. (1989). Perception of complex auditory patterns by humans. In S.H. Hulse and P. Dooling (Eds.) *The Comparative Psychology of Complex Acoustic Perception*, Lawrence Erlbaum, Hillsdale, NJ.

Fallon, Susan M. (1989). Discriminability of bursts of reproducible noise. Ph.D. Dissertation, Indiana University.

Kewley-Port, D., and Atal, B. (1989). Perceptual differences between vowels located in a limited phonetic space, *J. Acoust. Soc. Am.*, *85*, 1726-1740.

Kidd, G. R. (1989) Articulatory-rate context effects in phoneme identification. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 736-748.

Manuscripts in Progress

Espinoza-Varas, B. & Watson, C.S. Relations between auditory capabilities and phoneme perception. (in preparation for *J. Acoust. Soc. Am.*).

Kewley-Port, D. (1989). Detection thresholds for isolated vowels. *J. Acoust. Soc. Am.* (submitted).

Kidd, G.R. & Watson, C.S. Detection of changes in relative frequency in tonal sequences. (in preparation for *J. Acoust. Soc. Am.*).

Kidd, G.R. & Watson, C.S. Perception of multidimensional complex sounds. (in preparation for *Perception & Psychophysics*).

Watson, C.S., & Kidd, G.R. Discriminability of changes in tonal sequences is limited by the proportion of the total pattern duration that is changed. (in preparation for *J. Acoust. Soc. Am.*).

Watson, C.S., Foyle, D.C., & Kidd, G.R. Limited processing capacity for auditory pattern discrimination. (in preparation for *J. Acoust. Soc. Am.*).

Oral Presentations

Berg, B.G. & Robinson, D.E. (1987). Multiple observations and internal noise. *J. Acoust. Soc. Am.*, 81, S33

Espinoza-Varas, B., and Watson, C.S. (1987). Informational limits in speech processing by normal and hearing impaired listeners. *J. Acoust. Soc. Am.*, *82*, S42.

Fallon, S.M. & Robinson, D.E. (1987). The effects of decorrelation on the discriminability of noise samples. *J. Acoust. Soc. Am.*, 81, S33.

Kidd, G.R., and Watson, C.S. (1987). Perception of multidimensional complex sounds. *J. Acoust. Soc. Am.*, Suppl. 1, *81*, S33.

Watson, C.S., Espinoza-Varas, B., and Kyle, M.S. (1987). Rate and number as limiting factors in the perception of sequences of syllables. *J. Acoust. Soc. Am.*, *82*, S42.

Watson, C.S., and Kidd, G.R. (1987). Proportional target-tone duration as a factor in the discriminability of tonal patterns. *J. Acoust. Soc. Am.*, Suppl 1, *82*, S40.

Espinoza-Varas, B., and Watson, C.S. (1988). Low communality between tests of auditory discrimination and of speech perception. *J. Acoust. Soc. Am.*, Suppl 1, *84*, S143.

Kidd, G.R., and Watson, C.S. (1988). Detection of frequency- and time-transposed auditory patterns. *J. Acoust. Soc. Am.*, Suppl 1, *84*, S141.

Fallon, S.M., & Robinson, D.E. (1989). Effects of a silent temporal interval on discriminability of bursts of reproducible noise. *J. Acoust. Soc. Am.*, Suppl 1, *86*, S122.

Kewley-Port, D. (1989). Detection thresholds for isolated vowels. *J. Acoust. Soc. Am.*, Suppl 1, *85*, S51.

Kidd, G. R., & Watson, C. S. (1989) Detection of relative-frequency changes in tonal sequences. *J. Acoust. Soc. Am.*, Suppl 1, *86*, S121.

Rickert, M.E., & Robinson, D.E. (1989). Effects of temporal position on the detectability of interaurally uncorrelated noise. *J. Acoust. Soc. Am.*, Suppl 1, *86*, S12.

Watson, C.S. (1989). A performance-oriented approach to the hearing of complex sounds: Nonspeech and speech. *J. Acoust. Soc. Am.*, Suppl 1, *86*, S111.

Watson, C. S., & Kidd, G. R. (1989) Proportional target-tone duration as a limit on pattern discriminability: multi-component targets. *J. Acoust. Soc. Am.*, Suppl 1, *86*, S23.